On the Time Average of the Autocorrelation Function in Hamiltonian Dynamics

Pavle Saksida and Tomaž Prosen

Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia

E-mail: `Pavle.Saksida@fmf.uni-lj.si`, `Tomaz.Prosen@fmf.uni-lj.si`

Abstract. Rigorous lower bound on the time-average of the autocorrelation function of an arbitrary L^1 observable is proven in terms of conserved quantities and ergodic decompositions of the Hamiltonian dynamics. Improvements with respect to the bounds given by Mazur and Suzuki are discussed.

PACS numbers: 02.30.-f, 05.45.-a

AMS classification scheme numbers: 37K05, 37A25, 37A60

1. Introduction

Understanding the physical implications of ergodic and mixing properties of Hamiltonian dynamical systems is one of the main issues of non-equilibrium statistical mechanics. It is evident that a complete integrability in the sense of Liouville, or existence of even a small (non-complete) set of non-trivial global constants of the motion, is sufficient for a Hamiltonian system to be non-ergodic and non-mixing. It has been suggested by Mazur [1], and later followed up by Suzuki [2], that the existence of conservation laws can also be connected to divergent transport coefficients expressed - due to a *linear response theory* of Green and Kubo (see, e.g. [3, 4]) - in terms of the integrated time auto-correlation of the current observable. Namely, Green-Kubo formula expresses the conductivity κ_A with respect to a certain current observable A (say, the energy-current for the thermal conductivity) as

$$\kappa_A = \beta \int_0^\infty \langle A^0, A^t \rangle_\beta \,\mathrm{d}t,$$

where β is the inverse temperature, A^t denotes the time evolution of A. More precise meaning of the notation will be explained later.

Mazur has pointed out an inequality between the time-integrated or time-averaged auto-correlation of an arbitrary observable (which may also be interpreted thermodynamically as an *isothermal susceptibility*), and an algebraic expression which depends solely on the overlaps between the observable in question and the conservation laws expressed in terms of phase space integrals. His theorem is essentially a straightforward consequence of an ergodic theorem of Khinchine [5] and demonstrates that completely integrable systems should generically behave as ideal (ballistic) conductors of heat, electricity, etc.

Nevertheless, as such a result might seem quite natural from the point of view of dynamical systems and ergodic theory, it raised a lot of surprise and attention in the solid state community [6, 7]. Namely, as expressed in the language of solid state physics, one-dimensional completely integrable many-particle systems (with a large number of, or in the thermodynamic limit, an infinite number of degrees of freedom) generically posses a finite Drude weight, thus a divergent zero-frequency (d.c.) conductivity, and therefore behave as *ballistic conductors*, at *all temperatures*. Ballistic transport in a strongly interacting system at a non-vanishing temperature is certainly a statement which raises eyebrows of a solid state physicist. However, complete integrability is not even necessary to have such a striking physical implication, it is enough that at least one non-trivial conservation law exists which overlaps with the current [8].

In this paper we present some results which provide sharp and easily computable bounds of the autocorrelation function of an observable in terms of the conserved quantities of the system. Mazur's theorem can be considered as a linear (or first order) case of our bounds. Under certain conditions, which are quite likely to be fulfilled in many concrete cases, our results actually yield precise values of the time average of the

autocorrelation function. Mazur's proof relies heavily on a rather involved probabilistic and measure-theoretic statement of Khinchine [5]. Our method of proof is completely different, rather simpler and it avoids deep results of Khinchine entirely. Nevertheless, this method seems to give a good handle on the problem. In particular, it explains how the ergodic properties of the system on the one hand, and the form of the observable in question on the other, affect the sharpness of the bounds. We expect that saturability of our bounds can be an interesting test (probe) of complete integrability. We intend to address this issue in another paper.

In order to have geometric concepts well defined, we shall assume that the number of degrees of freedom is finite. Our results are essentially valid for the systems with the infinite number of degrees of freedom, which could be argued by taking the thermodynamic limit.

Our setup is the following. Let (M, ω, H) be a Hamiltonian system, where M is the phase space with the symplectic structure ω and the Hamiltonian H. Let

$$A : M \longrightarrow \mathbb{R}$$

be an arbitrary observable. By A^t we denote its time evolution given by

$$A^t(m) = A(\gamma_m(t)) : M \longrightarrow \mathbb{R},$$

where $\gamma_m(t)\colon \mathbb{R} \to M$ is the solution of the system (M, ω, H) satisfying the initial condition $\gamma_m(0) = m$.

The object of our interest in this paper will be the time average of the autocorrelation function

$$\mathcal{C}(A) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \langle A^0, A^t \rangle_\beta \; \mathrm{d}t.$$

The symbol $\langle -, - \rangle_\beta$ denotes the inner product on the L^2- space $L^2_\beta(M)$ given by

$$\langle F, G \rangle_\beta = \frac{1}{\mathcal{Z}(\beta)} \int_M F(m) \cdot G(m) \; e^{-\beta H(m)} \; \mathrm{d}m,$$

and

$$\mathcal{Z}(\beta) = \int_M e^{-\beta H(m)} \; dm$$

is the partition function which we shall assume to exist for strictly positive values of β (that is, for all finite temperatures if the system belongs to the realm of statistical physics). The measure $\mathrm{d}m$ is given by the top exterior power ω^r of the symplectic form, where r is the number of degrees of freedom. In local canonical coordinates:

$$\mathrm{d}m = \mathrm{d}q_1 \cdots \mathrm{d}q_r \mathrm{d}p_1 \cdots \mathrm{d}p_r.$$

The purpose of this paper is to give simple geometric descriptions of $\mathcal{C}(A)$ in terms of a set of, say k, functionally independent conserved quantities

$$H = H_1, H_2, \ldots, H_k : M \longrightarrow \mathbb{R}.$$

Our main result is the following bound on the time average of the autocorrelation function:

$$\mathcal{C}(A) \geq \sum_{\substack{\mathbf{l}=(l_1,\ldots,l_k),\ 0\leq|\mathbf{l}|\leq d \\ \mathbf{n}=(n_1,\ldots,n_k),\ 0\leq|\mathbf{n}|\leq d}} \langle A, H_1^{l_1}\ldots H_k^{l_k}\rangle_\beta \ (\mathbf{H}^{-1})_{\mathbf{l},\mathbf{n}} \ \langle A, H_1^{n_1}\ldots H_k^{n_k}\rangle_\beta, \tag{1}$$

where $|\mathbf{n}| = \sum_{i=1}^{k} n_i$. The elements of the matrix $\mathbf{H}$ are given by inner products

$$\mathbf{H}_{\mathbf{l},\mathbf{n}} = \langle H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k} \ , \ H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k}\rangle_\beta$$

in a suitably chosen order. Immediately after proving the above bound, we identify the conditions under which the bound becomes an equality. We note that the order d can be infinite.

Note that the strict positivity of $\mathcal{C}(A)$, implied by the strict positivity of the right-hand-side of (1), implies ballistic transport, if A is the corresponding current observable in the Green-Kubo theory.

The inequality (1) is a non-linear improvement of the inequality

$$\mathcal{C}(A) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \langle A^0, A^t\rangle_\beta \ \mathrm{d}t \ \geq \ \sum_{i,j=1}^{k} \langle A, H_i\rangle_\beta \cdot (\mathbf{H}^{-1})_{i,j} \cdot \langle A, H_j\rangle_\beta, \tag{2}$$

originally given by P. Mazur in [1]. Above $\mathbf{H}$ denotes the matrix with entries

$$\mathbf{H}_{i,j} = \langle H_i, H_j\rangle_\beta \tag{3}$$

and $\mathbf{H}^{-1}$ is its inverse.

In [1], working in the framework of statistical mechanics, Mazur treats the evolution $t \mapsto A(\gamma_m(t)) = A^t$ as a stochastic process. The main tool he uses is the power spectrum $I(\omega)$ of the process A^t. The power spectrum is given by the Fourier transform of the correlation function

$$\phi(t) = \langle A^0, A^t\rangle_\beta,$$

that is by

$$I(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t) \ e^{-it\omega} \ \mathrm{d}t.$$

The essential ingredient of Mazur's proof is the well-known result of Khinchine ([5]) which states that for every function $B\colon M \to \mathbb{R}$ we have

$$\mathcal{C}(B) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \langle A^0, A^t\rangle_\beta = I(0+) - I(0-) \geq 0.$$

From the above, Mazur's inequality follows almost immediately. The analytical crux of the matter is indeed contained in the above deep result of Khinchine.

Our treatment relies on direct and simple geometric considerations and does not resort to the theory of stochastic processes, let alone to the Khinchine's result. Despite its simplicity, our geometric approach enables us to get a better grip on the internality. We give simple and meaningful conditions for the inequality to be saturated. We also show how the ergodic properties of the system (M, ω, H) affect the quantity $\mathcal{C}(A)$.

Throughout this paper the observables $A\colon M \to \mathbb{R}$ will be assumed to be elements of $L^1_\beta(M)$, where in general $L^p_\beta(M)$ is the L^p-space generated by the measurable real functions $f\colon M \to \mathbb{R}$ which satisfy the integrability condition

$$\int_M |f|^p \, e^{-\beta H(m)} \mathrm{d}m \ < \infty.$$

Assuming that the partition function $\mathcal{Z}(\beta)$ is defined, the measure $e^{-\beta H(m)}\mathrm{d}m$ is finite. Therefore, we have the inclusion $L^1_\beta(M) \subset L^2_\beta(M)$, and thus A will also be an element of $L^2_\beta(M)$.

We will divide the discussion into two parts. Firstly, in section 2 we shall study the so-called ergodically regular case. This means that the dynamics is ergodic on the joint level sets of the conserved quantities. There the bound (1) is proven in theorem 2, whereas in theorem 1 we give a useful expression of $\mathcal{C}(A)$ as the L_2-norm of a certain projection of the observable A. Theorem 3 identifies the situation in which the bound (1) is saturated. It is reasonable to expect that the result of the theorem 3 will be useful in the context of the algebraically integrable systems, where the conserved quantities are expressible as polynomials or analytic functions. We list some sufficient conditions for the saturation in the remarks following theorem 3. Secondly, in section 3 we study the general case, without the assumption of ergodic regularity. The central technical result here is lemma 1, which is needed to prove the bound (1), (see corollary 2). In theorem 4 we prove that $\mathcal{C}(A)$ is equal to the L_2-norm of the orbital average of the observable A. In proposition 2 we show how ergodic decompositions can be used to further improve the bound (1).

2. The ergodically regular case

Let, as above, (M, ω, H) be a Hamiltonian system with $k < 2r = \dim(M)$ *functionally independent* conserved quantities

$$H = H_1, H_2, \ldots, H_k : M \longrightarrow \mathbb{R}$$

which are not necessarily in involution. Let

$$\begin{aligned} \mathcal{H} = (H_1, \ldots, H_k) :\ M &\longrightarrow D \subset \mathbb{R}^k \\ m &\longmapsto \mathcal{H}(m) = (\alpha_1, \ldots, \alpha_k) \end{aligned} \tag{4}$$

denote the moment map. We shall first consider the systems whose ergodic behaviour is simple in the sense that it is completely determined by the conserved quantities.

For the sake of brevity we shall denote the level sets of the moment map by

$$\mathcal{L}_\alpha = \mathcal{H}^{-1}(\alpha).$$

Definition 1 *The system (M, ω, H) is called ergodically regular, if it is ergodic on level sets $\mathcal{L}_\alpha$ for almost every $\alpha \in D$. More precisely, for almost every $\alpha \in D$ and for almost every $m \in \mathcal{L}_\alpha$ we have*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T B(\gamma_m(t)) \ \mathrm{d}t = \frac{1}{\mathrm{Vol}(\mathcal{L}_\alpha)} \int_{\mathcal{L}_\alpha} B(m_\alpha) \ \mathrm{d}m_\alpha,$$

where $B \in L^1_\beta(M)$ is an arbitrary observable, and $\mathrm{d}m_\alpha$ is the generalized microcanonical measure on $\mathcal{L}_\alpha$ induced by the measure $\mathrm{d}m$ which is given by ω^r.

A general example of ergodically regular systems are Liouville integrable systems. For the definition of the induced microcanonical measure, see [9]. In the above definition it is assumed that every volume $\mathrm{Vol}(\mathcal{L}_\alpha)$ is finite, but a moment of thought shows that this is a consequence of the finiteness of the partition function $\mathcal{Z}(\beta)$.

The key ingredient in the study of ergodically regular systems is the averaging map

$$B \stackrel{\mathcal{T}}{\longmapsto} B^{\mathcal{H}} = \frac{1}{\mathrm{Vol}(\mathcal{L}_\alpha)} \int_{\mathcal{L}_\alpha} B(h_\alpha)\, \mathrm{d}h_\alpha. \tag{5}$$

Let $L^1_\beta(D)$ be the L^1-space of measurable functions on D with respect to the measure μ_D given by

$$\mathrm{d}\mu_D = \frac{1}{\mathcal{Z}(\beta)}\, \mathrm{Vol}(\mathcal{L}_\alpha)\, e^{-\beta H^{\mathcal{H}}(\alpha)}\, \mathrm{d}\alpha = \frac{1}{\mathcal{Z}(\beta)}\, \mathrm{Vol}(\mathcal{L}_\alpha)\, e^{-\beta\alpha_1}\, \mathrm{d}\alpha. \tag{6}$$

Then the assignment (5) defines the operator

$$\begin{aligned} \mathcal{T}:\ L^1_\beta(M) &\longrightarrow L^1_\beta(D) \\ B &\longmapsto B^{\mathcal{H}}. \end{aligned} \tag{7}$$

Since our measures on M and D are both finite, we have $L^2_\beta(M) \subset L^1_\beta(M)$ and $L^2_\beta(D) \subset L^1_\beta(D)$. Obviously, the averaging operator descends to the map

$$\mathcal{T} : L^2_\beta(M) \longrightarrow L^2_\beta(D)$$

between the L^2-spaces. Let now the operator

$$\mathcal{R} : L^i_\beta(D) \longrightarrow L^i_\beta(M), \quad i = 1, 2$$

be given by

$$(\mathcal{R}(G))(m) = G(\mathcal{H}(m)).$$

The composed operator

$$\mathcal{P} = \mathcal{R} \circ \mathcal{T} : L^2_\beta(M) \longrightarrow L^2_\beta(M) \tag{8}$$

is then clearly a projector.

Proposition 1 *The operator $\mathcal{P}$ given by (8) is a continuous orthogonal operator and therefore symmetric on $L^2_\beta(M)$.*

Proof. To prove the orthogonality, let F be an element of the kernel of $\mathcal{P}$, and let G lie in its image. Then

$$\begin{aligned} \int_M F(m)\, G(m)\, e^{-\beta H(m)}\, \mathrm{d}m &= \int_D \mathrm{d}\alpha \int_{\mathcal{L}_\alpha} F(m_\alpha)\, G(m_\alpha)\, e^{-\beta H(m_\alpha)}\, \mathrm{d}m_\alpha \\ &= \int_D G^{\mathcal{H}}(\alpha)\, e^{-\beta H^{\mathcal{H}}(\alpha)} \Big(\int_{\mathcal{L}_\alpha} F(m_\alpha)\, \mathrm{d}m_\alpha \Big)\, \mathrm{d}\alpha \\ &= 0. \end{aligned}$$

Considering suitable convergent sequences in $L^2_\beta(M)$ it is easily seen that the subspaces $\mathrm{Ker}(\mathcal{P})$ and $\mathrm{Im}(\mathcal{P})$ are closed, which implies that $\mathcal{P}$ is a continuous orthogonal projection and therefore a symmetric map of $L^2_\beta(M)$ into itself.

□

From the above proposition we immediately obtain the basic geometric description of the time average $\mathcal{C}(A)$ in the ergodically regular case. By $\| - \|_D$ we shall denote the norm on $L^2_\beta(D)$ with respect to the measure $\mathrm{d}\mu_D$ introduced above.

Theorem 1 *Let A be an arbitrary measurable and integrable observable of the Hamiltonian system (M, ω, H). Then for the time average of its correlation function we have*

$$\mathcal{C}(A) = \|\mathcal{P}(A)\|^2_\beta = \|A^{\mathcal{H}}\|^2_D,$$

where $\| - \|_D$ denotes the L_2-norm on $L^2_\beta(D)$ given by the measure (6).

Proof. The fact that our system is ergodically regular gives

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T} \int_0^T \langle A^0, A^t \rangle_\beta \, \mathrm{d}t &= \lim_{T\to\infty} \frac{1}{T} \int_0^T \frac{1}{\mathcal{Z}(\beta)} \int_M A(\gamma_m(0)) \cdot A(\gamma_m(t)) \cdot e^{-\beta H(m)} \, \mathrm{d}m \\
&= \frac{1}{\mathcal{Z}(\beta)} \int_M \Big(A(\gamma_m(0)) \cdot \lim_{T\to\infty} \frac{1}{T} \int_0^T A(\gamma_m(t)) \;\; \cdot e^{-\beta H(m)} \; dt \Big) \, \mathrm{d}m \\
&= \frac{1}{\mathcal{Z}(\beta)} \int_M \Big(A(m) \cdot \frac{1}{\mathrm{Vol}(\mathcal{L}_{\mathcal{H}(\mathrm{m})})} \int_{\mathcal{L}_{\mathcal{H}(m)}} A(h) \, \mathrm{d}h \; \Big) \; e^{-\beta H(m)} \, \mathrm{d}m \\
&= \frac{1}{\mathcal{Z}(\beta)} \int_M A(m) \; \mathcal{P}(A)(m) \; e^{-\beta H(m)} \, \mathrm{d}m.
\end{aligned}$$

Using our notation and the fact that $\mathcal{P}$ is a symmetric projection we get the first equality of the theorem,

$$\langle A, \mathcal{P}(A) \rangle_\beta = \langle A, \mathcal{P}^2(A) \rangle_\beta = \langle \mathcal{P}(A), \mathcal{P}(A) \rangle_\beta = \|\mathcal{P}(A)\|^2_\beta.$$

Finally, for every pair of functions $A, B \colon M \to \mathbb{R}$ we have

$$\begin{aligned}
\langle \mathcal{P}(A), \mathcal{P}(B) \rangle_\beta &= \frac{1}{\mathcal{Z}(\beta)} \int_M \mathcal{P}(A)(m) \; \mathcal{P}(B)(m) \; e^{-\beta H(m)} \, \mathrm{d}m \\
&= \frac{1}{\mathcal{Z}(\beta)} \int_D \mathrm{d}\alpha \int_{\mathcal{L}_\alpha} \mathcal{P}(A)(m_\alpha) \; \mathcal{P}(B)(m_\alpha) \; e^{-\beta H(m_\alpha)} \, \mathrm{d}m_\alpha \qquad (9) \\
&= \frac{1}{\mathcal{Z}(\beta)} \int_D A^{\mathcal{H}}(\alpha) \; B^{\mathcal{H}}(\alpha) \; e^{-\beta\alpha_1} \mathrm{Vol}(\mathcal{L}_\alpha) \, \mathrm{d}\alpha \\
&= \langle A^{\mathcal{H}}, B^{\mathcal{H}} \rangle_D,
\end{aligned}$$

which proves our second equality.

□

We shall now turn to the analogues of the right-hand side of Mazur's inequality. These expressions will give us estimates for $\mathcal{C}(A)$. In concrete cases these expressions are likely to be simpler to calculate than the norm of $A^{\mathcal{H}}$. More importantly, they will be useful in the treatment of the general, ergodically irregular case.

Let us consider the set of all monomials corresponding to the multi-indices $\mathbf{n} = (n_1, \ldots, n_k)$ of degree d or less

$$H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k}, \qquad 0 \leq |\mathbf{n}| = \ n_1 + n_2 + \ldots n_k \leq d$$

composed of the conserved quantities $H = H_1, H_2, \ldots, H_k$ of our system. We shall order these monomials by a combination of the ordering by degree and the lexicographical ordering,

$$\textit{if } |\mathbf{n}_1| < |\mathbf{n}_2|, \textit{ then } \mathbf{n}_1 < \mathbf{n}_2 \ ; \quad \textit{if } |\mathbf{n}_1| = |\mathbf{n}_2|, \textit{ then ordered lexicographically.} \tag{10}$$

By V_d we shall denote the subspace of $L^2_\beta(M)$ spanned by the above monomials,

$$V_d = \mathrm{span}\{H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k};\ |\mathbf{n}| \leq d\} \subset L^2_\beta(M).$$

The fact that the conserved quantities are functionally independent implies the linear independence of the monomials. The basis $\{H_1^{n_1} \ldots H_k^{n_k}\}$ of V_d is not orthonormal, therefore we will need the matrix of the inner product $\langle -, - \rangle_\beta$ on V_d corresponding to our basis. The elements of the inner product matrix $\mathbf{H}$ are given by

$$\begin{aligned}(\mathbf{H})_{o(\mathbf{l}),o(\mathbf{n})} &= \langle H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k}\ ,\ H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k} \rangle_\beta \\ &= \int_M H_1^{l_1+n_1}(m) H_2^{l_2+n_2}(m) \ldots H_k^{l_k+n_k}(m)\ e^{-\beta H(m)}\ \mathrm{d}m,\end{aligned}$$

where $o(\mathbf{l})$, $o(\mathbf{n})$ are integers given by the ordering (10). The matrix $\mathbf{H}$ is non-singular due to the linear independence of the monomials.

Theorem 2 *Let A be an observable on an ergodically regular Hamiltonian system (M, ω, H) with the additional conserved quantities $H_2, \ldots H_k$. Then for every positive integer d we have*

$$\mathcal{C}(A) \geq \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\nu(d)} \langle A, H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k} \rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k} \rangle_\beta, \tag{11}$$

where $\nu(d)$ denotes the number of different monomials of degrees ranging between 0 *and d in k unknowns.*

The bound (11) could be of practical importance, since it is relatively easily calculable in many cases. We notice that the left-hand side of (1) is independent of d, therefore theorem 2 has the following immediate corollary.

Corollary 1 *Let (M, ω, H) and $A\colon M \to \mathbb{R}$ be as above. Then*

$$\mathcal{C}(A) \geq \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\infty} \langle A, H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k} \rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k} \rangle_\beta. \tag{12}$$

□

Clearly, the bound (12) is sharp. From the practical point of view, it is in general less useful. The main problem is the evaluation of the inverse of the infinite matrix $\mathbf{H}$. This can be tackled by replacing the monomials by products of suitably scaled orthogonal polynomials.

Proof of Theorem 2. Let $\mathcal{U} = \mathcal{P}(L^2_\beta(M)) \subset L^2_\beta(M)$ be the image of the projector $\mathcal{P}$. This is a closed subspace in $L^2_\beta(M)$ and the inherited inner product gives it the structure of a Hilbert space. Let

$$\pi : \mathcal{U} \longrightarrow V_d$$

be the orthogonal projection. We shall prove that the expression on the right-hand side of (11) is equal to the norm of the vector $\pi(\mathcal{P}(A))$. Since by Pythagoras' theorem any orthogonal projection of a vector is shorter than the vector itself, (11) will follow immediately from theorem 1.

Let us introduce a shorter notation

$$h_{o(\mathbf{n})} = H_1^{n_1} H_2^{n_2} \dots H_k^{n_k}$$

for the elements of the basis of V_d, and let $\{h_i^*; i = \nu(d)+1, \nu(d)+2, \dots\}$ be the basis of the orthogonal complement $V_d^\perp \subset \mathcal{U}$ composed of the vectors from the basis dual to $\{h_n\}_{n\in\mathbb{N}}$. This means

$$\langle h_i^*, h_j\rangle_\beta = 0, \quad for \quad i = 0 \dots, \nu(d), \quad j = \nu(d)+1, \nu(d)+2, \dots .$$

Let

$$\mathcal{P}(A) = \sum_{i=0}^{\nu(d)} c_i\ h_i\ +\ \sum_{j=\nu(d)+1}^{\infty} d_j\ h_j^* \tag{13}$$

be the orthogonal decomposition of $\mathcal{P}(A)$ with respect to $\mathcal{U} = V_d \oplus V_d^\perp$. Taking the inner product of this expression with each of h_k for $k = 1, \dots, \nu(d)$ gives

$$\langle \mathcal{P}(A), h_k\rangle_\beta = \sum_{i=0}^{\nu(d)} c_i \langle h_i, h_k\rangle_\beta .$$

Inverting the matrix $\mathbf{H}_{i,k} = \langle h_i, h_k\rangle_\beta$ yields the expression of the coefficients c_i:

$$c_i = \sum_{j=0}^{\nu(d)} (\mathbf{H}^{-1})_{i,j}\ \langle \mathcal{P}(A), h_j\rangle_\beta . \tag{14}$$

The inner product matrix of $V_d \subset L^2_\beta(M)$ with respect to the basis $\{h_i, \dots, h_{\nu(d)}\}$ is $\mathbf{H}$, thus the squared norm $\|\pi(\mathcal{P}(A)\|^2_\beta$ of the orthogonal projection of $\mathcal{P}(A)$ onto V_d is equal to

$$\|\pi(\mathcal{P}(A)\|^2_\beta = (c_0, \dots, c_{\nu(d)}) \cdot \mathbf{H} \cdot \begin{pmatrix} c_0 \\ \vdots \\ c_{\nu(d)} \end{pmatrix} .$$

By means of (14) we can express this quantity in terms of the products $\langle \mathcal{P}(A), h_i\rangle_\beta$. The inner product matrix $\mathbf{H}$ is replaced by $\mathbf{H}^{-1} \cdot \mathbf{H} \cdot \mathbf{H}^{-1}$, so we get

$$\|\pi(\mathcal{P}(A))\|^2 = (\langle \mathcal{P}(A), h_0\rangle_\beta, \dots, \langle \mathcal{P}(A), h_{\nu(d)}\rangle_\beta) \cdot \mathbf{H}^{-1} \cdot \begin{pmatrix} \langle \mathcal{P}(A), h_0\rangle_\beta \\ \vdots \\ \langle \mathcal{P}(A), h_{\nu(d)}\rangle_\beta \end{pmatrix} .$$

To complete the proof we only have to show that $\langle A, h_i\rangle_\beta = \langle \mathcal{P}(A), h_i\rangle_\beta$ for every observable A on M. But, the monomials $h_{o(\mathbf{n})} = H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k}$ are constant on every level set $\mathcal{L}_\alpha$, and so

$$\begin{aligned}
\langle A, h_i\rangle_\beta &= \int_M A(m) h_i(m)\ e^{-\beta H(m)}\ \mathrm{d}m \\
&= \int_D h_i^{\mathcal{H}}(\alpha)\ e^{-\beta H^{\mathcal{H}}(\alpha)} \Big(\int_{\mathcal{L}_\alpha} A(m_\alpha)\ \mathrm{d}m_\alpha \Big)\ \mathrm{d}\alpha \qquad (15)\\
&= \int_D A^{\mathcal{H}}(\alpha) h_i^{\mathcal{H}}(\alpha)\ e^{-\beta\alpha_1}\ \mathrm{d}\alpha \\
&= \langle A^{\mathcal{H}}, h_i^{\mathcal{H}}\rangle_D .
\end{aligned}$$

From (10) and from the fact that $\mathcal{P}(h_i) = h_i$, we now finally get

$$\langle A^{\mathcal{H}}, h_i^{\mathcal{H}}\rangle_D = \langle \mathcal{P}(A), \mathcal{P}(h_i)\rangle_\beta = \langle \mathcal{P}(A), h_i\rangle_\beta = \langle \mathcal{P}(A), H_1^{m_1} H_2^{m_2} \ldots H_k^{m_k}\rangle_\beta,$$

which concludes the proof.

□

We shall now clarify the question, when the inequalities in (11) and in (12) are saturated. The answer is given by the theorem 3 bellow, which is more or less an immediate corollary of theorem 2.

Theorem 3 **(A)** *Suppose an observable A on M satisfies the two equivalent conditions:*

(i) The function $A^{\mathcal{H}}\colon D \to \mathbb{R}$ is a polynomial of degree d in the variables $(\alpha_1, \ldots, \alpha_k)$.

(ii) The observable A be expressible in the form

$$A(m) = \sum_{o(\mathbf{n})=0}^{\nu(d)} c_{o(\mathbf{n})}(m)\ H_1^{n_1}(m) \cdots H_k^{n_k}(m), \qquad (16)$$

where

$$c^{\mathcal{H}}_{o(\mathbf{n})}(\alpha) = \frac{1}{\mathrm{Vol}(\mathcal{L}_\alpha)} \int_{\mathcal{L}_\alpha} c_{o(\mathbf{n})}(m_\alpha)\ \mathrm{d}m_\alpha \equiv const. \quad for\ every\ \mathbf{n}. \qquad (17)$$

Then

$$\mathcal{C}(A) = \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\nu(d)} \langle A, H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k}\rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k}\rangle_\beta. \qquad (18)$$

(B) *If $A^{\mathcal{H}}\colon D \to \mathbb{R}$ is an* analytic *function, or alternatively, if A is expressible in the form*

$$A(m) = \sum_{o(\mathbf{n})=0}^{\infty} c_{o(\mathbf{n})}(m)\ H_1^{n_1}(m) \cdots H_k^{n_k}(m),$$

where the coefficients $c_{o(\mathbf{n})}$ again satisfy the condition (17), then we have the equality

$$\mathcal{C}(A) = \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\infty} \langle A, H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k}\rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k}\rangle_\beta. \qquad (19)$$

Proof. First we check the fact that the conditions *1.* and *2.* of the theorem are indeed equivalent. Let the observable A be expressible as

$$A(m) = \sum_{o(\mathbf{n})=0}^{\nu(d')} d_{o(\mathbf{n})}(m) H_1^{n_1}(m) \cdots H_k^{n_k}(m).$$

For every m, such that $\mathcal{H}(m) = (\alpha_1, \dots, \alpha_k)$, we have $H_i(m) \equiv \alpha_i$ on $\mathcal{L}_\alpha$. Therefore,

$$A^{\mathcal{H}}(\alpha) = \sum_{o(\mathbf{n})=0}^{\nu(d')} d^{\mathcal{H}}_{o(\mathbf{n})}(\alpha)\ \alpha_1^{n_1} \cdots \alpha_k^{n_k}.$$

This function is a polynomial precisely when all $d^{\mathcal{H}}_{o(\mathbf{n})}(\alpha)$ are polynomials. In such cases the above function can be rewritten in the form

$$A^{\mathcal{H}}(\alpha) = \sum_{o(\mathbf{n})=0}^{\nu(d)} \widetilde{c}_{o(\mathbf{n})}\ \alpha_1^{n_1} \cdots \alpha_k^{n_k},$$

where the constants $\widetilde{c}_{o(\mathbf{n})}$ are coefficients of the polynomials $d^{\mathcal{H}}_{o(\mathbf{n})}(\alpha)$. Clearly, every $d_{o(\mathbf{n})}$ is of the form $d_{o(\mathbf{n})}(m) = \sum c_j(m) H_1^{r_1}(m) \cdots H_k^{r_k}(m)$ for some choice of the multi-indices $(r_1, \dots r_k)$, and for every $c_j(m)$ we have $c_j^{\mathcal{H}} = \widetilde{c}_j$, which proves the equivalence of *1.* and *2.*

To establish the validity of **(A)**, let

$$A^{\mathcal{H}}(\alpha_1, \dots, \alpha_k) = \sum_{o(\mathbf{n})=0}^{\nu(d)} \widetilde{c}_{o(\mathbf{n})}\ \alpha_1^{n_1} \alpha_2^{n_2} \dots \alpha_k^{n_k}$$

be a polynomial. Then the pull-back of $A^{\mathcal{H}}$ on M is given by

$$(\mathcal{L}(A^{\mathcal{H}}))(m) = (\mathcal{P}(A))(m) = \sum_{o(\mathbf{n})=0}^{\nu(d)} \left(c_{o(\mathbf{n})}\ H_1^{n_1} H_2^{n_2} \dots H_k^{n_k} \right)(m) = \sum_{o(\mathbf{n})=0}^{\nu(d)} c_{o(\mathbf{n})}(m) h_{o(\mathbf{n})}(m).$$

According to (13) this means that $\mathcal{P}(A)$ lies in the subspace V_d of the space $L^2_\beta(M)$, and thus

$$\pi(\mathcal{P}(A)) = \mathcal{P}(A).$$

As seen above, we then have

$$\begin{aligned} \mathcal{C}(A) &= \|\mathcal{P}(A)\|^2_\beta = \|\pi(\mathcal{P}(A))\|^2_\beta \\ &= \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\nu(d)} \langle A, H_1^{l_1} H_2^{l_2} \dots H_k^{l_k} \rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \dots H_k^{n_k} \rangle_\beta . \end{aligned}$$

Part **(B)** of the theorem is an easy consequence of part **(A)**.

□

Bellow we collect some remarks and comments that illustrate the condition (17).

Remarks 1. *The coefficients $c_{o(\mathbf{n})}(m)$ in the expression (16) are, of course, in general not constants. They can be rather arbitrary functions that change in the fibre direction of the moment map $\mathcal{H}\colon M \to D$, as well as in the directions transversal to the fibres $\mathcal{L}_\alpha$.*

2. Let $\alpha \in D \subset \mathbb{R}^k$ be a regular value of the moment map $\mathcal{H}: M \to D$. Then there exists a neighbourhood $\alpha \in U \subset D$ such that the open subset $\mathcal{H}^{-1}(U) \subset M$ is diffeomorphic to $U \times \mathcal{L}$ and $\mathcal{L}$ is diffeomorphic to $\mathcal{L}_\alpha$. Open sets $W \subset \mathcal{H}^{-1}(U) \cong U \times \mathcal{L}$ can be coordinatized as $W = \{(\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l)\}$, where $(\alpha_1, \ldots, \alpha_k)$ are coordinates on $U \subset D$ and $(\beta_1, \ldots \beta_l)$ some local coordinates on a patch of $\mathcal{L}$. Then the conditions

$$\frac{\partial}{\partial \alpha_i} c_{o(\mathbf{n})} \equiv 0, \qquad i = 1, \ldots, k, \quad o(\mathbf{n}) = 0, \ldots, \nu(d) \tag{20}$$

are sufficient for (17) and hence for (18). Clearly, the above conditions are not necessary for (17). The functions $c_{o(\mathbf{n})}$ are allowed to vary in the α-directions, only their averages over $\mathcal{L}_\alpha$ have to be constant.

3. Let M be an almost Kähler manifold. This means that it is equipped with a metric $g(-,-)$ compatible with the symplectic form in the sense that there exists an almost complex structure J on M such that

$$g_m(X_m, Y_m) = \omega_m(X_m, J_m(Y_m)), \qquad X_m, Y_m \in T_m M.$$

The most common examples of such manifolds are cotangent bundles $M = T^*N$ over Riemannian manifolds N. The metric g_N on N is extended in the natural way onto the metric on the tangent bundle TN, and the symplectic form ω^T on TN is given as the exterior derivative $\omega^T = \mathrm{d}\theta$ of the tautological 1-form

$$\theta_m(v_m) = (g_N)_{\pi(m)}\Big((D_m\pi)(v_m), m\Big), \qquad v_m \in T_m(TN).$$

Here $\pi: TN \to N$ is the natural projection. (Note that m is a tangent vector, $m \in T_{\pi(m)}N$.) One can use the metric again to pull the form ω^T back to the cotangent bundle T^*N. Then we have the well defined gradient vector fields ∇H_i associated to functions H_i, which are g-orthogonal to the Hamiltonian vector fields X_H. Moreover, the gradients ∇H_i are orthogonal to the entire fibres $\mathcal{L}_\alpha$. Therefore, the conditions

$$(\nabla H_i)(c_{o(\mathbf{n})}) \equiv 0, \quad i = 1, \ldots, k, \quad o(\mathbf{n}) = 0, \ldots, \nu(d)$$

are equivalent to the conditions (20). Similarly as the conditions (20), they are sufficient but not necessary for (17).

4. The condition that $A^{\mathcal{H}}$ is a polynomial is also equivalent to the condition

$$\textit{if } |\mathbf{n}| = n_1 + n_2 + \ldots + n_k > d, \qquad \textit{then} \quad \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k} \rangle_\beta = 0.$$

This follows immediately from the expansion (13) and from the equivalence of the conditions 1. and 2. of theorem 3. In practice, the problem with the above condition is that infinitely many integrals have to be checked. But on the other hand, in some cases, the integrals of the functions $A \cdot H_1^{n_1} \cdots H_k^{n_k}$ might be rather easily computable in some concrete cases.

3. The general case

The ergodically regular systems treated above are very special. The invariant subspaces on which a Hamiltonian system is ergodic can in general be rather wild, and even within

a single system they can be of very different types, e.g. full level sets $\mathcal{L}_\alpha$, subsets of $\mathcal{L}_\alpha$ of lower dimensionality, as well as subsets of $\mathcal{L}_\alpha$ having even fractal dimensions. One should recall for example typical situations of smooth perturbations of integrable systems landing in the context of the Kolmogorov-Arnold-Moser theory. In general, it is impossible to parameterize invariant ergodic sets by some manageable (say Hausdorff) space over which one could integrate. For these reasons we have to modify our approach in order to prove our bounds for a general Hamiltonian system. In particular, for a useful description of the ergodic decomposition of our system, we shall revert to an inverse limit type construction. This will enable us to prove the results analogous to those from the previous section, but valid for the general Hamiltonian systems. The analogue of theorem 1 is the following:

Theorem 4 *Let (M, ω, H) be an arbitrary Hamiltonian system with a well defined partition function and let $A\colon M \to \mathbb{R}$ be an element of $L^1_\beta(M)$. Then for the orbital average*

$$\widetilde{A}(m) = \lim_{T\to\infty} \frac{1}{T} \int_0^T A(\gamma_m(t))\ \mathrm{d}t$$

we have

$$\mathcal{C}(A) = \|\widetilde{A}\|^2_\beta.$$

Before giving the proof, we shall describe the ergodic decomposition of the Hamiltonian system (M, ω, H) which will be used in the proofs. In our construction we shall use the invariant measure on M given by

$$\mu(N) = \frac{1}{\mathcal{Z}(\beta)} \int_N e^{-\beta H(m)}\ \mathrm{d}m.$$

As in the previous section, the crux of the proof will be the replacement of the temporal averages by the spatial ones in the context without ergodic regularity. To be able to do this, we have to decompose the space M into a collection of invariant sets on which our system *is* ergodic. We will construct such ergodic decomposition by means of successive approximations. By definition the Hamiltonian system (M, ω, H) is *not* ergodic, if there exists an invariant measurable set $N \subset M$ such that

$$0 < \mu(N) < 1,$$

where the inequalities have to be strict. Let the first approximation of our ergodic decomposition be a finite partition $\mathcal{N}_1$ consisting of *invariant sets*

$$N^1_1, N^1_2, \ldots, N^1_{k_1} \subset M$$

with the properties

$$0 < \mu(N^1_i) < 1, \quad \mu(N^1_i \cap N^1_j) = 0, \quad \cup_{i=1}^{k_1} N^1_i = M.$$

In the next stage we decompose each N^1_i into *invariant* measurable sets $\{N^2_{k_i}, N^2_{k_i+1}, \ldots, N^2_{l_i}\}$ with analogous properties. This yields the partition $\mathcal{N}_2 =$

$\{N^2_1, \dots N^2_{k_2}\}$ which again satisfies the stipulations analogous to those listed above. We continue the procedure and obtain a sequence $\{\mathcal{N}_n\}_{n\in\mathbb{N}}$ of partitions in which every term $\mathcal{N}_n = \{N^n_1, \dots, N^n_{k_n}\}$ satisfies

$$0 < \mu(N^n_i) < 1, \quad \mu(N^n_i \cap N^n_j) = 0, \quad \cup_{i=1}^{k_n} = M.$$

Let now $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of natural numbers such that for every $i \in \mathbb{N}$ we have

$$1 \le a_i \le k_i.$$

Then the set

$$N_{\{a_n\}} = \lim_{n\to\infty} N^n_{a_n} = \cup_{n=1}^{\infty} N^n_{a_n}$$

is measurable and invariant. Moreover, the system (M, ω, H) is ergodic on every $N_{\{a_n\}}$.
Remark *The sets $N_{\{a_n\}}$ are subsets of level sets $\mathcal{L}_\alpha$, therefore they are sets of measure zero. These sets can be strange, of noninteger Hausdorff dimensions etc... The collection of $N_{\{a_n\}}$ is parameterized by some subset of the set of sequences $\{a_n\}_{n\in\mathbb{N}}$ of the form described above. In order to simplify the parameter set, one could make the above construction in a "binary manner" by decomposing each set N^n_i into only two invariant sets N^{n+1}_0 and N^{n+1}_1. The sequences $\{a_n\}_{n\in\mathbb{N}}$ would then be the maps*

$$\{a_n\}_{n\in\mathbb{N}} : \mathbb{N} \longrightarrow \{0, 1\}.$$

Now we construct the sequence $\{A_n\}_{n\in\mathbb{N}}$ of measurable functions

$$A_n : M \longrightarrow \mathbb{R}$$

by the rule

$$if \quad m \in N^n_i \quad then \quad A_n(m) = \frac{1}{\mathrm{Vol}(N^n_i)} \int_{N^n_i} A(h)\ \mathrm{d}h = C_{n,i}. \tag{21}$$

Proof of Theorem 4: Our strategy here will be analogous to the one used in the proof of theorem 1. If $A\colon M \to \mathbb{R}$ is an element of $L^1_\beta(M)$, then by Birkhoff's theorem (see, e. g. [10]) the function $\widetilde{A}$ is also measurable and $\widetilde{A} \in L^1_\beta(M)$. We clearly have

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \ \widetilde{A}(\gamma_m(t))\ \mathrm{d}t = \widetilde{A}(m),$$

which means that the operator

$$\begin{aligned} \mathcal{A} :\ \ L^2_\beta(M) \ &\longrightarrow \ L^2_\beta(M) \\ A \ &\longmapsto \ \widetilde{A} \end{aligned}$$

is a projector. The image of $\mathcal{A}$ are the those functions in $L^2_\beta(M)$ which are constant on every H-orbit. This is a closed subspace in $L^2_\beta(M)$, therefore $\mathcal{A}$ is a continuous operator. Moreover, it is also an *orthogonal* projector. To see this, let $B \in \ker\mathcal{A}$, and $\widetilde{A} = \mathcal{A}(A)$ be arbitrary elements in the kernel and in the image of $\mathcal{A}$ respectively. We claim that

$$\langle B, \widetilde{A}\rangle_\beta = \int_M B(m)\ \widetilde{A}(m)\ \mathrm{d}\mu = 0. \tag{22}$$

Let $\{A_n\}_{n\in\mathbb{N}}$ be the sequence of functions approximating A as described above. By construction we have: For every $m \in M$ there exists a unique sequence $\{a_n\}_{n\in\mathbb{N}}$ such that $m \in N_{\{a_n\}} = \lim_{n\to\infty} N^n_{a_n}$. Therefore,

$$\widetilde{A}(m) = \lim_{n\to\infty} A_n(m),$$

since our system is ergodic on every $N_{\{a_n\}}$. This implies

$$B(m)\widetilde{A}(m) = \lim_{n\to\infty} B(m)A_n(m).$$

Functions A and B are elements of L^1_β. Let now Sup be the essential supremum of the orbital average $\widetilde{A}$ on M. The measure μ on M is finite, so a function Sup is an element of $L^1_\beta(M)$. We have the inequality

$$B(m)A_n(m) \leq B(m)\ Sup$$

which holds for almost every $m \in M$. Therefore, by the Lebesgue dominated convergence theorem we have

$$\int_M B(m)\ \widetilde{A}(m)\ \mathrm{d}\mu = \lim_{n\to\infty} \int_M B(m)\ A_n(m)\ \mathrm{d}\mu. \tag{23}$$

Since A_n takes the constant value $C_{i,n}$ on every N^n_i for $i = 1, \ldots, k_n$, and since all the sets N^n_i as well as the measure μ are H-invariant, the Liouville theorem gives

$$\begin{aligned}\int_M B(m)\ A_n(m)\ \mathrm{d}\mu &= \sum_{i=1}^{k_n} C_{i,n} \int_{N^n_i} B(m)\ \mathrm{d}\mu \\ &= \sum_{i=1}^{k_n} C_{i,n} \int_{N^n_i} B(\gamma_m(t))\ \mathrm{d}\mu \end{aligned} \tag{24}$$

for every $t \in \mathbb{R}$. Since $\mathcal{A}(B) = \widetilde{B} = 0$, we get

$$\int_M B(m)\ A_n(m)\ \mathrm{d}\mu = \sum_{i=1}^{k_n} C_{i,n} \lim_{T\to\infty} \frac{1}{T} \int_0^T \Big(\int_{N^n_i} B(\gamma_m(t))\ \mathrm{d}\mu \Big)\ \mathrm{d}t = 0,$$

which together with (23) proves (22).

Now, every orthogonal projection is a symmetric operator, therefore

$$\begin{aligned}\mathcal{C}(A) &= \lim_{T\to\infty} \frac{1}{T} \Big(\int_M A(m)\ A(\gamma_m(t))\ \mathrm{d}\mu \Big)\ \mathrm{d}t \\ &= \langle A, \mathcal{A}(A)\rangle_\beta = \langle A, \mathcal{A}^2(A)\rangle_\beta \\ &= \langle \mathcal{A}(A), \mathcal{A}(A)\rangle_\beta \\ &= \|\widetilde{A}\|^2_\beta,\end{aligned}$$

which completes the proof.

□

In most cases the orbital average $\widetilde{A}$ is impossible to calculate. Therefore, theorem 4 almost never provides a good estimate for $\mathcal{C}(A)$, apart from the fact that it ensures $\mathcal{C}(A)$ to be nonnegative. Therefore, as we have done in the ergodically regular case, we shall construct more easily calculable estimates which will use the specific information

about the Hamiltonian system in question. In the general case this information has two sources, the conserved quantities and some ergodic decomposition $\{\mathcal{N}_n\}_{n\in\mathbb{N}}$ of the form described above.

In the proof of theorem 4 we used the sequence $\{A_n\}_{n\in\mathbb{N}}$ of functions approximating the observable A. We shall now modify the approximating sequence, so that it will take into account the conserved quantities $H_1,\ldots,H_k$ of the system as well as the ergodic decomposition. We shall mimic the approach from the previous section, but will replace the averaging over the level sets $\mathcal{L}_\alpha$ with the averaging over the intersections

$$N^n_{\alpha,i} = \mathcal{L}_\alpha \cap N^n_i.$$

Consider the functions $B_{n,i}\colon M \to \mathbb{R}$ for $n \in \mathbb{N}$ and $i \in \{1,\ldots k_n\}$ associated to the ergodic decomposition $\{\mathcal{N}_n\}_{n\in\mathbb{N}}$, and given by the rule

$$B_{n,i}(m) = \begin{cases} \frac{1}{\mathrm{Vol}(N^n_{\mathcal{H}(m),i})}\int_{N^n_{\mathcal{H}(m),i}} A(m_\alpha)\;\mathrm{d}m_\alpha & ; \quad m \in N^n_i \\ \\ 0 & ; \quad \textit{otherwise} \end{cases}. \tag{25}$$

Let now the sequence $\{C_n\}_{n\in\mathbb{N}}$ of functions

$$C_n : M \longrightarrow \mathbb{R}$$

be given by

$$C_n(m) = \sum_{i=1}^{k_n} B_{n,i}(m). \tag{26}$$

The functions C_n are approximations of the orbital average of A. In addition, their behaviour is similar to that of the projection $\mathcal{P}(A)$ in the ergodically regular case. More precisely, the functions C_n are refinements of $\mathcal{P}(A)$ which take into account the partition $\mathcal{N}_n$.

Recall that

$$(\mathcal{P}(A))(m) = \frac{1}{\mathrm{Vol}(\mathcal{L}_{\mathcal{H}(m)})}\int_{\mathcal{L}_{\mathcal{H}(m)}} A(m_\alpha)\;\mathrm{d}m_\alpha, \quad \textit{for } m \in \mathcal{L}_\alpha.$$

Since $\mathcal{L}_\alpha = \cup_{i=1}^{k_n} N^n_{\alpha,i}$, for every $m \in \mathcal{L}_\alpha$, we have

$$\begin{aligned}(\mathcal{P}(A))(m) &= \frac{1}{\mathrm{Vol}(\mathcal{L}_\alpha)}\sum_{i=1}^{k_n}\int_{N^n_{\alpha,i}} A(m_\alpha)\;\mathrm{d}m_\alpha \\ &= \sum_{i=1}^{k_n} C_n(m_i)\,\frac{\mathrm{Vol}(N^n_{\alpha,i})}{\mathrm{Vol}(\mathcal{L}_\alpha)} \qquad m_i \in N^n_{\mathcal{H}(m),i}\ \textit{arbitrary}.\end{aligned} \tag{27}$$

Lemma 1 *The sequence of functions $\{C_n\}_{n\in\mathbb{N}}$ has the following properties.*

(i) For every $n \in \mathbb{N}$ we have

$$\mathcal{C}(C_n) = \|C_n\|^2_\beta \geq \|\mathcal{P}(A)\|^2_\beta = \|A^{\mathcal{H}}\|^2_D.$$

(ii) The sequence $\{\|C_n\|^2_\beta\}_{n\in\mathbb{N}}$ is non-decreasing.

(iii) The sequence $\{\|C_n\|_\beta^2\}_{n\in\mathbb{N}}$ *is convergent, and*

$$\mathcal{C}(A) = \lim_{n\to\infty} \mathcal{C}(C_n) = \lim_{n\to\infty} \|C_n\|_\beta^2.$$

Proof. *Ad 1.* The definition of the time average of the correlation function, and the fact that the functions C_n are constant on the H-invariant sets $N_{\alpha,i}^n$ give

$$\begin{aligned}
\mathcal{C}(C_n) &= \int_M C_n(m)\Big(\lim_{T\to\infty}\frac{1}{T}\int_0^T C_n(\gamma_m(t))\ \mathrm{d}t\Big)\ \mathrm{d}\mu \\
&= \int_D \mathrm{d}\alpha \sum_{i=1}^{k_n}\int_{N_{\alpha,i}} B_{n,i}(m_\alpha)\Big(\lim_{T\to\infty}\frac{1}{T}\int_0^T B_{n,i}(\gamma_{m_\alpha}(t))\ \mathrm{d}t\Big)\ \mathrm{d}m_\alpha \\
&= \int_D \mathrm{d}\alpha \sum_{i=1}^{k_n}\int_{N_{\alpha,i}^n} B_{n,i}^2(m_\alpha)\ \mathrm{d}m_\alpha.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\|C_n\|_\beta^2 &= \int_M C_n(m)^2\ \mathrm{d}\mu \\
&= \int_D\ \mathrm{d}\alpha \sum_{i=1}^{k_n}\int_{N_{\alpha,i}^n} B_{n,i}^2(m_\alpha)\ \mathrm{d}m_\alpha, \qquad (28)
\end{aligned}$$

which proves the first equality in *1.*

To prove the inequality $\|C_n\|_\beta^2 \geq \|\mathcal{P}(A)\|_\beta^2$ we use (27) and the usual procedure for calculating variances. This gives

$$\begin{aligned}
0 &\leq \sum_{i=1}^{k_n}\Big(C_n(m_i) - \mathcal{P}(A)(m)\Big)^2\ \frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)} \\
&= \sum_{i=1}^{k_n} C_n^2(m_i)\frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)} - \mathcal{P}(A)^2(m) \qquad m_i \in N_{\alpha,i}^n\ \textit{ arbitrary}
\end{aligned}$$

for every $m \in \mathcal{L}_\alpha$. Integrating the above inequality over M with respect to the measure $\mathrm{d}\mu$ gives

$$\int_M\Big(\sum_{i=1}^{k_n} C_n^2(m_i)\frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)}\Big)\ \mathrm{d}\mu\ \geq\ \int_M \mathcal{P}(A)^2(m)\ \mathrm{d}\mu = \|\mathcal{P}(A)\|_\beta^2.$$

Observing that the expression $\sum_{i=1}^{k_n} C_n^2(m_i)\frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)}$ is an average of a function over $\mathcal{L}_\alpha$ and is therefore constant on $\mathcal{L}_\alpha$ yields

$$\begin{aligned}
\int_M\Big(\sum_{i=1}^{k_n} C_n^2(m_i)\frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)}\Big)\ \mathrm{d}\mu &= \int_D \mathrm{d}\alpha\int_{\mathcal{L}_\alpha}\sum_{i=1}^{k_n} C_n^2(m_i)\frac{\mathrm{Vol}(N_{\alpha,i}^n)}{\mathrm{Vol}(\mathcal{L}_\alpha)}\ \mathrm{d}m_\alpha \\
&= \int_D\Big(\sum_{i=1}^{k_n} C_n^2(m_i)\mathrm{Vol}(N_{\alpha,i}^n)\Big)\ \mathrm{d}\alpha.
\end{aligned}$$

Since $C_n(m_i) = B_{n,i}(m)$ for suitable pairs of m and m_i, we see from (28) that the above expression is indeed equal to the norm $\|C_n\|_\beta^2$, which concludes the proof of the point *1.*

Ad 2. The proof that for every $n \in \mathbb{N}$ we have $\|C_{n+1}\|_\beta^2 \geq \|C_n\|_\beta^2$, is essentially the same as the proof of the inequality $\|C_n\|_\beta^2 \geq \|\mathcal{P}(A)\|_\beta^2$ just given. Above we partitioned the phase space into the disjoint union $\cup_{i=1}^{k_n} N_i^n = M$ of H-invariant subsets with positive measures. To prove *2.* we have to partition every N_i^n further into the union $\cup_{j=l_i}^{k_i} N_j^{(n+1)} = N_i^n$. The actual calculations here are then precisely the same as in *1.*, modulo slightly different notation.

Ad 3. First we observe that for every $m \in M$ we have

$$\lim_{n\to\infty} C_n(m) = \widetilde{A}(m).$$

Indeed, for every m there exists a unique α such that $m \in \mathcal{L}_\alpha$, and a unique sequence $\{a_n\}_{n\in\mathbb{N}}$ such that $m \in N_{\{a_n\}} = \cap_{n=1}^\infty N_{a_n}^n$. Since our system is ergodic on the limit set $N_{\{a_n\}}^n = \lim_{n\to\infty} N_{a_n}^n$, and since the orbit $\gamma_m(t)$ is contained in $\mathcal{L}_\alpha$, we have

$$\begin{aligned}
\lim_{n\to\infty} C_n(m) &= \lim_{n\to\infty} \frac{1}{\mathrm{Vol}(N_{\alpha,a_n}^n)} \int_{N_{\alpha,a_n}^n} A(m_\alpha)\ \mathrm{d}m_\alpha \\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T A(\gamma_m(t))\ \mathrm{d}t \\
&= \widetilde{A}(m).
\end{aligned}$$

Let again *Sup* be the essential supremum of the orbital average $\widetilde{A}$ on M. Then clearly

$$C_n(m) \leq Sup; \quad \textit{a. e. on } M,$$

since the values $C(m)$ are averages taken over larger sets than those that yield $\widetilde{A}(m)$. As we already mentioned, *Sup* is an element of $L_\beta^1(M)$. Thus, the sequence $\{C_n^2\}_{n\in\mathbb{N}}$ together with its point-wise limit $\widetilde{A}^2$ satisfies the conditions of the Lebesgue dominated convergence theorem. Therefore, we have

$$\lim_{n\to\infty} \|C_n\|_\beta^2 = \lim_{n\to\infty} \int_M C_n^2(m)\ \mathrm{d}\mu = \int_M \widetilde{A}^2(m)\ \mathrm{d}\mu = \|\widetilde{A}\|_\beta^2.$$

This, together with the result of theorem 4, concludes the proof.

□

An immediate consequence of lemma 1 is the fact that all the inequalities proved in the previous section for the ergodically regular systems hold for all Hamiltonian systems without restrictions. In particular, we have

Corollary 2 *Let (M, ω, H) be an arbitrary Hamiltonian system with k conserved quantities $H = H_1, H_2, \ldots H_k$. Then for every observable $A: M \to \mathbb{R}$ and every $d \in \mathbb{N}$ we have:*

$$\mathcal{C}(A) \geq \sum_{o(\mathbf{l}),o(\mathbf{n})=0}^{\nu(d)} \langle A, H_1^{l_1} H_2^{l_2} \ldots H_k^{l_k} \rangle_\beta\ (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{n})}\ \langle A, H_1^{n_1} H_2^{n_2} \ldots H_k^{n_k} \rangle_\beta. \tag{29}$$

The above inequality also holds for $d = \infty$.

Proof. We have just proved that for a general system we have $\mathcal{C}(A) \geq \mathcal{P}(A)$. Proposition then follows from theorem 2.

□

We notice that the estimate in the above proposition does not reflect in any way how far our system is from being ergodic on the level sets $\mathcal{L}_\alpha$. The non-ergodicity is reflected in the ergodic decomposition $\{\mathcal{N}_n\}_{n\in\mathbb{N}}$, and the information given by $\{\mathcal{N}_n\}_{n\in\mathbb{N}}$ can be used to improve the bound (29). In our setup it is quite easy to plug the decomposition $\mathcal{N}_n$ into (29). Recall the definitions (25) and (26). Since for $i \neq j$ the supports N_i^n and N_j^n of the functions $B_{n,i}$ and $B_{n,j}$ are disjoint, we have

$$\langle B_{n,i}, B_{n,j}\rangle_\beta = 0. \tag{30}$$

The function $C_n = \sum_{i=1}^{k_n} B_{n,i}$ is a sum of orthogonal vectors, therefore,

$$\|C_n\|_\beta^2 = \sum_{i=1}^{k_n} \|B_{n,i}\|_\beta^2. \tag{31}$$

As in the previous section, we can project orthogonally the function $C_n \in L^2_\beta(M)$ on the subspace $V_d \subset L^2_\beta(M)$ spanned by the monomials $H_1^{n_1}H_2^{n_2}\dots H_k^{n_k}$ of degree d or less. We have proved that the vector C_n is longer than $\mathcal{P}(A)$, therefore for the projections by $\pi\colon L^2_d(M) \to V_d$ we have

$$\pi(C_n) \geq \pi(\mathcal{P}(A)). \tag{32}$$

Taking into account (30) and (31), and lemma 1, we get the following proposition.

Proposition 2 *For every $n \in \mathbb{N}$ we have*

$$\mathcal{C}(A) \geq \sum_{i=1}^{k_n}\Bigg(\sum_{o(\mathbf{l}),o(\mathbf{j})=0}^{\nu(d)} \langle B_{n,i}, H_1^{l_1}H_2^{l_2}\dots H_k^{l_k}\rangle_\beta \cdot (\mathbf{H}^{-1})_{o(\mathbf{l}),o(\mathbf{j})} \cdot \langle B_{n,i}, H_1^{j_1}H_2^{j_2}\dots H_k^{j_k}\rangle_\beta\Bigg).$$

The inequality also holds for $d = \infty$.

□

The quality of the above estimate increases with increasing n and d, and for every $n \geq 2$ the above estimate is better that the estimate in theorem 2.

Acknowledgements

The authors would like to express their gratitude to Mirko Degli Esposti, Andreas Knauf and Oliver Dragičević for their valuable comments and suggestions. The work has been supported by the programmes P1-0291, P1-0044, and the grant J1-7347 of the Slovenian Research Agency (ARRS).

References

[1] Mazur P 1969 Physica **43** 533

[2] Suzuki M 1971 Physica **51** 277

[3] Kubo R 1957 J. Phys. Soc. Japan **12** 570

[4] For a very clear account on Green-Kubo theory in the realm of classical mechanics see: Visscher W M 1974 Phys. Rev. A **10** 2461

[5] Khinchine A I 1934 Ann. Math. **109** 604
[6] Zotos X, Naef F and Prelovšek P 1997 Phys. Rev. B **55** 11029
[7] Zotos X and Prelovšek P, review article published in "Strong Interactions in Low Dimensions", series "Physics and Chemistry of Materials with Low Dimensional Structures", eds. D. Baeriswyl and L. Degiorgi, Kluwer Academic Publishers (2004)
[8] Prosen T 1998 J. Phys. A: Math. Gen. **31** L645
[9] Minlos R A 2002 *Introduction to Mathematical Statistical Physics* (In Russian) Moscow: MCNMO)
[10] Mane R 1987 *Ergodic Theory and Differentiable Dynamics* (Berlin: Springer-Verlag)